

## A9. Aprendizaje Autosupervisado



**XUNTA DE GALICIA**

CONSELLERÍA DE CULTURA,  
EDUCACIÓN E UNIVERSIDADE

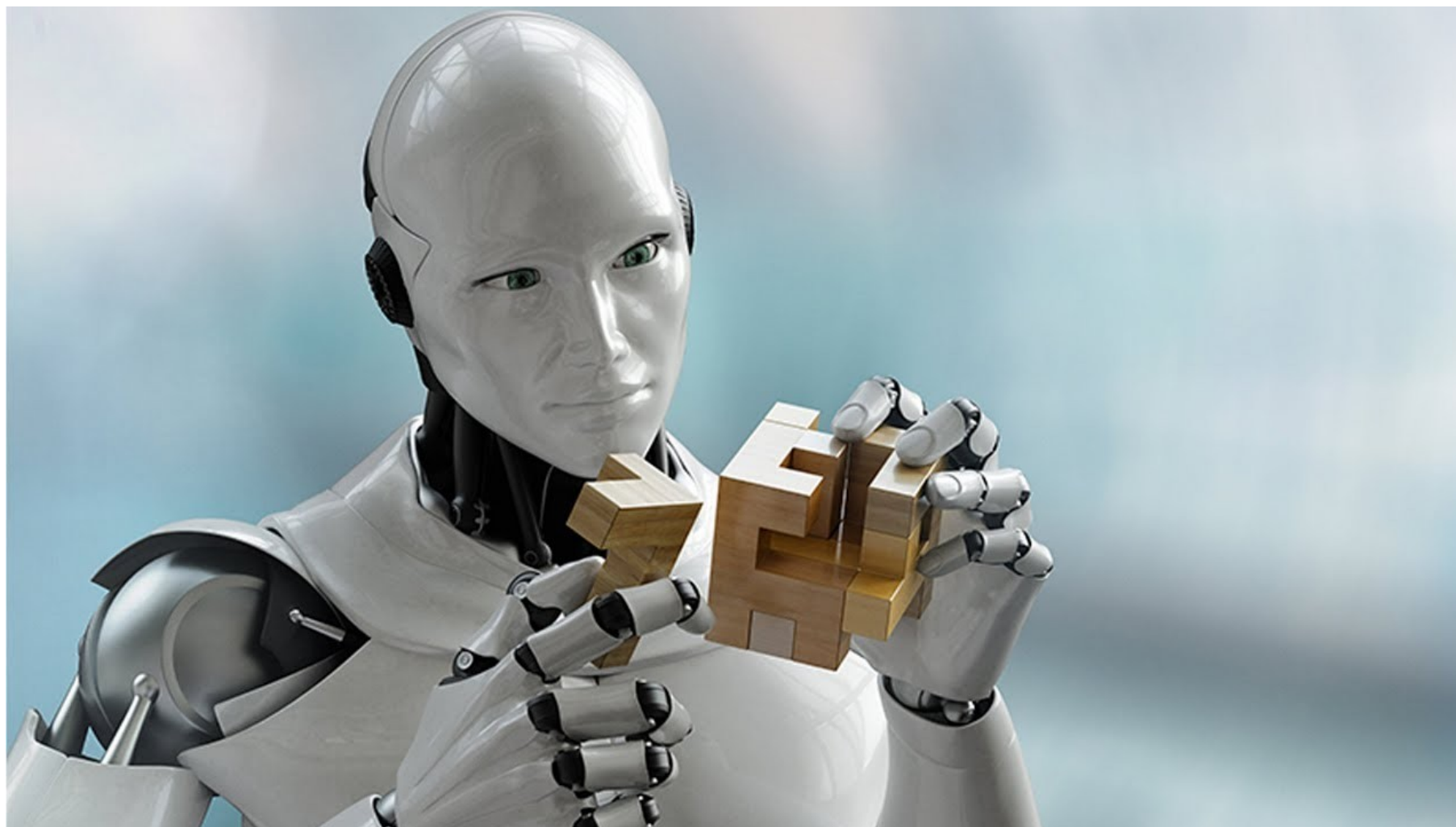


**IES de Teis**

Avda. de Galicia, 101  
36216 – Vigo  
886 12 04 64  
[ies.teis@edu.xunta.es](mailto:ies.teis@edu.xunta.es)



*Unión Europea-  
NextGenerationEU*



## A9. Aprendizaje Autosupervisado

### Índice.

1. ¿Qué es el Aprendizaje Auto Supervisado?.....	3
2. Categorías importantes.....	5
3. Ventajas.....	6
4. Desventajas.....	7
5. Limitaciones.....	8
6. Beneficios.....	9
7. Técnicas.....	10
8. ¿Para qué sirve el Aprendizaje Auto Supervisado?.....	11
9. Ejemplos de Aprendizaje Auto Supervisado.....	12
10. Evolución del Aprendizaje.....	13

## A9. Aprendizaje Autosupervisado

### 1. ¿Qué es el Aprendizaje Auto Supervisado?

---



## A9. Aprendizaje Autosupervisado

### 1. ¿Qué es el Aprendizaje Auto Supervisado?

---

El **Aprendizaje Auto Supervisado**, es un enfoque en el campo del Aprendizaje Automático donde un modelo se entrena utilizando sus propias salidas como objetivos de aprendizaje; en lugar de depender de etiquetas externas, el modelo genera automáticamente tareas de entrenamiento a partir de datos disponibles.

Este enfoque busca descubrir patrones subyacentes y representaciones útiles en los datos sin necesidad de etiquetas externas, permitiendo que los modelos adquieran un entendimiento más profundo y generalizado de los datos, lo que es particularmente útil en escenarios donde las etiquetas son escasas o costosas.

El Aprendizaje Auto Supervisado ha tenido un gran éxito en:

- Robótica
- Procesamiento del Lenguaje Natural. Ej: modelo BERT de Google.

## A9. Aprendizaje Autosupervisado

### 2. Categorías importantes.

El Aprendizaje Auto Supervisado se divide en dos categorías principales:

- **Aprendizaje Preentrenado:** utiliza conjuntos de datos existentes para crear un modelo predictivo.

El Aprendizaje Autosupervisado Preentrenado funciona tomando un conjunto de datos etiquetados y entrenando un modelo con ellos.

El modelo se utiliza entonces para identificar patrones en los nuevos datos y hacer predicciones .

Este modelo se utiliza con conjuntos de datos demasiado grandes para ser etiquetados por personas, como imágenes para aplicaciones de visión por ordenador.



**Aprendizaje en línea:** actualiza continuamente sus modelos a medida que adquiere nuevos datos.

El Aprendizaje Autosupervisado en línea, también conocido como **Aprendizaje por Refuerzo**, es un tipo de Aprendizaje Automático en el que se entrena un sistema para que ajuste su comportamiento en función de la información que recibe de su entorno.

Este tipo de aprendizaje permite aprender a un sistema con el tiempo y adaptarse a nuevas situaciones y datos.



## A9. Aprendizaje Autosupervisado

### 3. Ventajas.

---

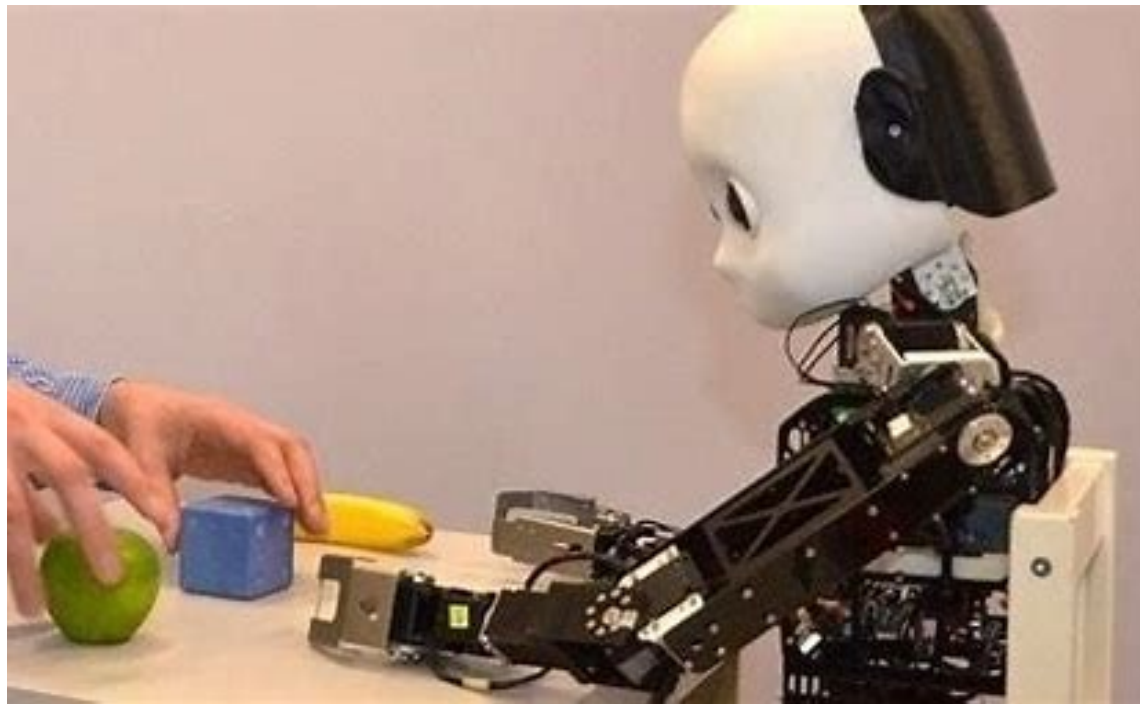
El **aprendizaje auto-supervisado** tiene las siguientes ventajas:

- Permitir a los modelos aprender a partir de datos no etiquetados, lo que reduce la necesidad de etiquetado manual y amplía los conjuntos de datos disponibles para el entrenamiento.
- Ayudar a descubrir patrones y características ocultas en los datos sin la necesidad de intervención humana.
- **Reducción de la dependencia de los datos etiquetados:** uno de los principales beneficios del aprendizaje autosupervisado es que permite que un modelo aprenda características y representaciones útiles de los datos sin la necesidad de grandes cantidades de datos etiquetados. Esto puede resultar especialmente útil en situaciones en las que obtener datos etiquetados resulta caro o lleva mucho tiempo, o en las que hay una cantidad limitada de datos etiquetados disponibles.
- **Mejora en la generalización:** el aprendizaje autosupervisado puede mejorar el rendimiento de la generalización de un modelo, lo que significa que puede realizar predicciones más precisas sobre datos invisibles. Esto se debe a que el aprendizaje autosupervisado permite que un modelo aprenda de la estructura inherente de los datos, en lugar de simplemente memorizar ejemplos específicos.
- **Transferencia de aprendizaje:** el aprendizaje autosupervisado puede resultar útil para el aprendizaje por transferencia, que implica el uso de un modelo entrenado en una tarea para mejorar el rendimiento en una tarea relacionada. Al aprender características y representaciones útiles de los datos mediante el aprendizaje autosupervisado, un modelo se puede adaptar más fácilmente a nuevas tareas y entornos.
- **Escalabilidad:** el aprendizaje autosupervisado puede ser más escalable que el aprendizaje supervisado, ya que permite que un modelo aprenda de un conjunto de datos más grande sin necesidad de anotaciones humanas. Esto puede resultar especialmente útil en situaciones en las que la cantidad de datos es demasiado grande para que los humanos puedan etiquetarlos.

### 4. Desventajas.

El **aprendizaje auto-supervisado** tiene las siguientes desventajas:

- Dificultad de evaluar objetivamente la calidad del aprendizaje debido a la falta de etiquetas de verdad absoluta.
- Puede requerir una mayor capacidad computacional y tiempo de entrenamiento en comparación con los enfoques supervisados tradicionales.



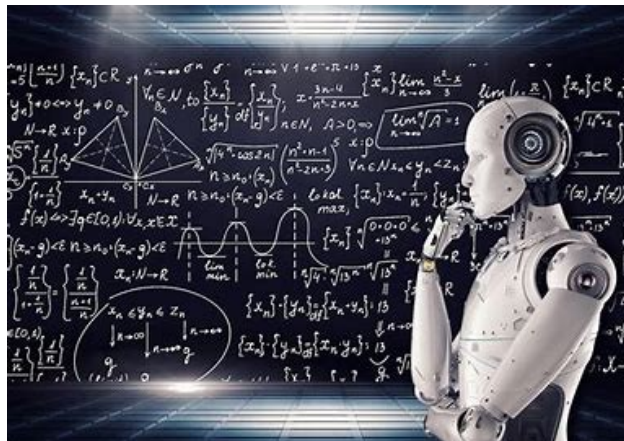


## A9. Aprendizaje Autosupervisado

### 5. Limitaciones.

El **aprendizaje auto-supervisado** tiene las siguientes limitaciones:

- Calidad de la señal de supervisión: puede llegar a ser ruidosa y conducir a un menor rendimiento.
- Limitado a ciertos tipos de tareas: puede no ser tan efectivo para tareas con datos muy complejos o no estructurados.
- Complejidad de la formación: algunas técnicas pueden ser más complejas de implementar y entrenar que las del aprendizaje supervisado.
- Necesidad de mayor potencia computacional en comparación con el aprendizaje supervisado.
- Se basa en un modelo para dar sentido a los datos no etiquetados y generar etiquetas.
- Evaluación del rendimiento de los modelos de aprendizaje puede ser desalentadora porque no hay un punto de referencia estándar para la comparación.





## A9. Aprendizaje Autosupervisado

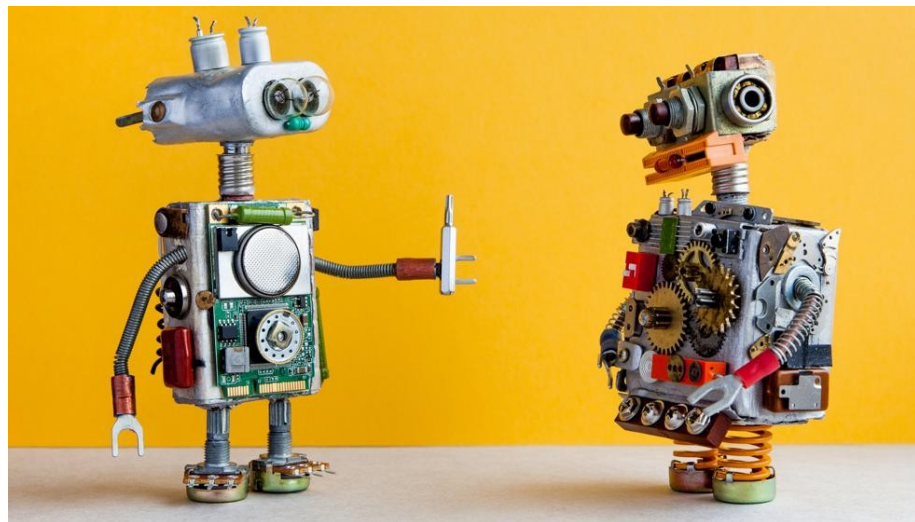
### 6. Beneficios.

---

El **aprendizaje auto-supervisado** ofrece toda una serie de beneficios al campo de la Inteligencia Artificial:

- Capacidad de pre-entrenar modelos en grandes conjuntos de datos sin la necesidad de etiquetas.
- Mejor capacidad de generalización en tareas específicas.
- Reducción de la dependencia de conjuntos de datos etiquetados, lo que permite un aprendizaje más eficiente y escalable.
- Facilita la transferencia de conocimiento entre dominios y tareas.
- Permite la exploración y el descubrimiento de características y relaciones ocultas en los datos

Estos beneficios hacen que el aprendizaje auto-supervisado sea un enfoque poderoso para impulsar el avance de la IA y su adopción en una variedad de industrias y aplicaciones.



## A9. Aprendizaje Autosupervisado

### 7. Técnicas.

---

El **Aprendizaje Autosupervisado** abarca una serie de técnicas que aprovechan los datos no etiquetados para permitir que los modelos aprendan y den sentido a patrones complejos de forma autónoma. Estas técnicas han revolucionado varios dominios, incluido el procesamiento del lenguaje natural (PLN) y la comprensión de videos

Algunas técnicas clave del **Aprendizaje Autosupervisado** son las siguientes:

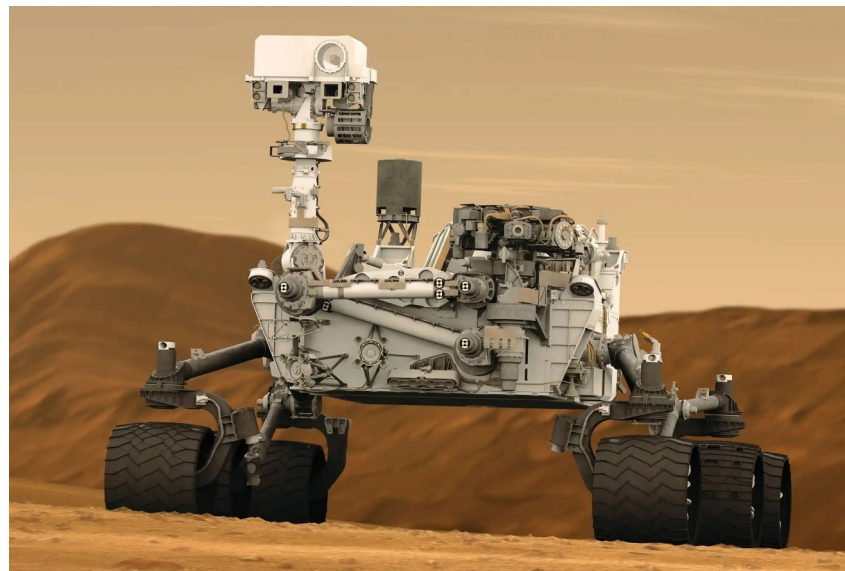
- **Modelado del lenguaje:** En el PLN, los modelos se entrenan para predecir las palabras que faltan en una oración, lo que les permite aprender la estructura subyacente y la semántica del lenguaje..
- **Modelado de lenguaje enmascarado:** esta técnica consiste en enmascarar aleatoriamente palabras en una oración y entrenar al modelo para predecir las palabras que faltan. Ayuda al modelo a comprender el contexto y el significado de las palabras.
- **Predicción de fotogramas de vídeo:** para la comprensión de vídeo, los modelos se entrenan para predecir el siguiente fotograma de una secuencia. Esto permite que el modelo aprenda las dependencias temporales y capture representaciones significativas de los vídeos.
- **Aprendizaje contrastivo:** Esta técnica entrena al modelo para diferenciar entre instancias similares y diferentes, lo que mejora su capacidad para capturar detalles y similitudes de grano fino.
- **Discriminación de instancias:** los modelos se entrenan para discriminar entre instancias de la misma clase, lo que promueve un aprendizaje sólido de características y mejora la generalización.

## A9. Aprendizaje Autosupervisado

### 8. ¿Para qué sirve el Aprendizaje Auto Supervisado?

El **aprendizaje auto supervisado** desempeña un papel esencial al abordar desafíos en situaciones donde la disponibilidad de etiquetas de entrenamiento es limitada. Al eliminar la necesidad de etiquetas externas, este enfoque permite que los modelos aprendan representaciones útiles y significativas de los datos a través de tareas de auto-generación. Esto resulta especialmente valioso en dominios donde la adquisición de etiquetas es costosa o impracticable, como en grandes conjuntos de datos no etiquetados. Además, también contribuye a la creación de modelos más generales y transferibles, ya que aprenden patrones subyacentes sin depender de anotaciones específicas.

Otra aplicación destacada de este tipo de modelo algorítmico se encuentra en la exploración y comprensión de características latentes en datos complejos. Al aprovechar estructuras inherentes en los datos, este enfoque facilita la identificación de relaciones y patrones, permitiendo a los modelos capturar información valiosa sin la guía explícita de etiquetas externas. Esto hace que el aprendizaje autosupervisado sea crucial en la mejora de la capacidad de generalización y adaptabilidad de los modelos, lo que es esencial cuando la diversidad y complejidad de los datos requieren enfoques flexibles y robustos.



### 9. Ejemplos del Aprendizaje Auto Supervisado.

---

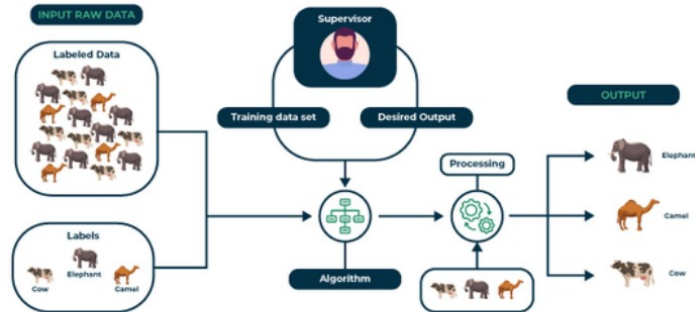
Algunos ejemplos del uso del **aprendizaje autosupervisado** son los siguientes:

- Wav2vec de Facebook para realizar un reconocimiento del habla a través de dos redes neuronales convolucionales profundas que se complementan entre sí.
- Modelo de Representación de Codificador Bidireccional de Transformadores, BERT (en inglés: Bidirectional Encoder Representations from Transformers) de Google se utiliza para comprender mejor el contexto de las consultas de búsqueda.
- GPT 3, GPT 4, ..., de Open AI, como modelo del lenguaje autorregresivo que se utiliza en el procesamiento del lenguaje, traducción de textos, respuesta a preguntas, etc.
- Bootstrap Your Own Latent es un NCSSL que produjo excelentes resultados en ImageNet y en pruebas comparativas de transferencia y semisupervisión.
- El algoritmo Yarowsky es un ejemplo de aprendizaje autosupervisado en el procesamiento del lenguaje natural. A partir de un pequeño número de ejemplos etiquetados, aprende a predecir qué sentido de una palabra polisémica se utiliza en un punto determinado del texto.
- DirectPred es un NCSSL que establece directamente los pesos del predictor en lugar de aprenderlo mediante la actualización del gradiente.
- Representaciones de imágenes.
- Embedding de texto.
- Análisis de series temporales.
- Reconocimiento de actividades humanas.
- Clasificación de audio.
- Identificación de patrones biológicos.

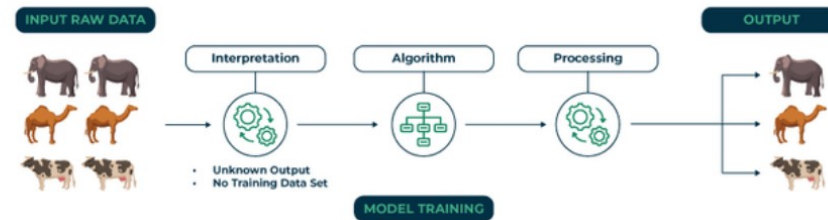
## A9. Aprendizaje Autosupervisado

### 10. Evolución del Aprendizaje.

Aprendizaje Supervisado



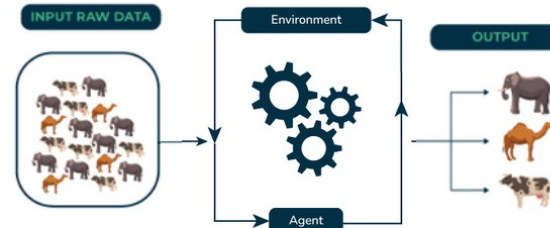
Aprendizaje No Supervisado



Aprendizaje Semi Supervisado



Aprendizaje Reforzado



Aprendizaje Auto Supervisado

